

Fernando BRONCANO

Resumen

En este trabajo nos proponemos defender una versión de la tesis de la “mente extendida” frente a las críticas de algunos autores para quienes los sistemas que llaman “extracraneales” no cumplirían las condiciones que deben cumplir los componentes de los procesos mentales. Proponemos la integración de funciones como un criterio para considerar un conjunto de procesos como un proceso mental. Desde este punto de vista, consideramos que, en ciertas situaciones, algunos sistemas externos podrían llegar a cumplir este criterio

Palabras clave: Mente extensa, subjetividad, procesos mentales, unidad de la mente.

Abstract

In this paper we aim to defend a version of the thesis of “extended mind” against the criticism of some authors that consider that the “extracraneal” devices cannot accomplish the requirements that the components of mental processes must meet. We propose a quality of integration as a criterion to be a mental process, and we consider that, in some situations, external devices can be considered as meeting this criterion.

Keywords: Extended mind, subjectivity, mental processes, unity of mind.

¹ Conversaciones con Jesús Vega, Miguel Ángel Salichs y Ricardo Sanz, Manuel Bedia y Liza Skidelsky, quien hizo sugerencias y críticas valiosas en el Encuentro Hispano-Argentino de Filosofía Analítica (Eduardo Rabossi *in memoriam*), han sido determinantes en algunas de las ideas de este trabajo. A todos mi reconocimiento y agradecimiento. Este trabajo se ha realizado en el marco del proyecto BFF2003-04372 2003-6 del MEC.

1. Fronteras de la mente y el cuerpo

De entre las consecuencias metafísicas del funcionalismo, quizá la más importante sea la invitación a una revisión de la identidad humana y personal. No es exagerado considerar que el siglo XX ha sido el siglo de la revolución metafísica del funcionalismo: entraña una novedosa partición de lo real que acoge nuevas clases de sistemas causales individualizados por su función y no por los procesos particulares que los constituyen. Estas nuevas clases incluyen todos los sistemas que describimos bajo la perspectiva de su diseño: artefactos, también los sistemas de procesamiento de información, seres vivos, y, especialmente, los propios sistemas componentes de la mente. Aplicado a la mente humana, el funcionalismo ha contribuido al desarrollo de las ciencias cognitivas y a algunas transformaciones importantes en la psicología. Hay numerosas formas de funcionalismo que sería largo siquiera enumerar, cuando no discutir, pero quisiera llamar la atención hacia algunas formas de que están transformando nuestra imagen como seres en el mundo, aunque todavía tentativamente y en forma exploratoria; modificando, quiero decir, nuestra ontología de sujetos y personas. Me refiero a una complicada clase que incluye entradas como “mente encarnada”, “mente distribuida”, “seres híbridos”, “transhumanismo”, ciborgs, etc. No quisiera por el momento preocuparme por si hay o no características consistentes que estén en la intersección de todos estos proyectos, más allá de que son síntomas de un rechazo a una concepción cartesiana dualista del ser humano. Se trata de exploraciones en nuevas formas de funcionalismo que se alejan del computacionalismo clásico enraizado en la metáfora de la mente como ordenador y recuperan temas y territorios heideggerianos como la situacionalidad, la apertura, etc., e incluso hegelianos como el reconocimiento y la lucha por el reconocimiento. En particular quisiera referirme a una de estas variedades del nuevo jardín metafísico que ha sido denominada por sus proponentes como *mente extendida* (Chalmers-Clark 1998). En un breve artículo de 1998 Chalmers y Clark propusieron esta concepción que ahora discutiremos con una intención abiertamente provocativa, que ha tenido cierto impacto aunque no muy grande en la filosofía sí mayor en las corrientes de diseño de sistemas artificiales, donde ha sido incorporado como un término más del argot técnico. Quizá ni siquiera sus autores le hayan dado demasiada importancia filosófica, mas precisamente por este breve alcance pretendido de la tesis me parece que es un término merecedor de una reflexión filosófica con algún interés antropológico.

El funcionalismo, como es sabido, tiene como una tesis nuclear el que no podemos considerar ya que ni el material, ni la forma y diseño como criterios suficientes para individuar un sistema funcional, sino que es la función o el conjunto de funciones características de ese sistema las que determinan la identidad ontológica. Si adoptamos esta perspectiva de manera muy liberal para incluir las funciones que lle-

van la “marca de lo mental”, es decir, aquéllas que portan contenidos simbólicos y son susceptibles de interpretación, el funcionalismo se convertiría en una forma muy abstracta de caracterizar lo mental que incluiría o podría incluir, convenientemente cualificado, a estas nuevas aproximaciones tan sugerentes y novedosas aplicadas en este caso a delimitar las fronteras existenciales de lo mental. ¿A qué llamamos mente extendida? : según sus autores, que toman el término por analogía con el fenotipo extendido con el que Richard Dawkins caracterizaba los artefactos y técnicas animales, una extensión de la mente es un sistema integrado que incorpora de modo esencial y no accidental sistemas de información y que es susceptible de convertirse en una prótesis para funciones específica y característicamente mentales . El principio básico es el que Chalmers y Clark, han denominado:

PRINCIPIO DE PARIDAD: Si al enfrentarnos a cierta tarea, una parte del mundo funciona como un proceso que si estuviese en la cabeza no dudáramos en aceptarlo como una parte del proceso cognitivo, entonces esa parte del mundo es (en ese momento) parte del proceso cognitivo

Veamos los dos pasos en los que se puede dividir este principio.

DESCARGA COMPUTACIONAL de datos y tareas en el medio. La mente humana es un sistema computacionalmente abierto, que se completa con la colaboración del medio, en el que descarga tantos datos y procesos computacionales como puede.

La mente humana, sostiene Dennett (1996, cap. 5) descarga en el medio tanta computación como puede: “Guardamos “punteros” e “índices” en nuestro cerebro y tenemos tanto datos como podemos en el mundo externo, en nuestros cuadernos de direcciones, bibliotecas, cuadernos de notas, computadoras –y, claro, también en nuestro círculo de amigos y asociados. Una mente humana que estuviera limitada al cerebro estaría gravemente discapacitada si le quitásemos esos instrumentos externos (...) Cuantos más datos y mecanismos descargamos, dependemos más de esos periféricos” (Dennett 1996, 144). La evolución de nuestra especie tiene mucho que ver con la transformación del medio producida por la construcción de estas herramientas cognitivas y artefactos epistémicos². Merlin Donald (1991) sostiene, en este sentido, también en que la especie humana se caracteriza por hacer del medio un almacén externo de símbolos. En otro trabajo (Broncano 2004) he defendido la coevolución de la mente humana, como sintonizadora natural de “semióforos” o entidades portadoras de significados, una mente inmersa en un mundo lleno de sig-

² El término se debe a Vega (2006). Es un trabajo recomendable y accesible para complementar este.

nos, que ha llegado a desarrollar la capacidad de construir símbolos a través de sus capacidades metarrepresentacionales. Estas capacidades técnicas de los humanos contribuyen a crear un medio, un entorno (para usar la vieja diferencia de Ortega) en el que la evolución personal y de especie se desarrolla condicionada por las interacciones con estos artefactos epistémicos, símbolos y semióforos sintonizados. Nuestras descargas computacionales tienen por ello el carácter de descargas de tareas informacionales en sistemas simbólicos.

El principio de descarga computacional es mucho más fácil de entender gráficamente que de explicar en términos técnicos: cuando deseamos conocer el teléfono de la compañía de gas buscamos en la guía telefónica, pero no nos aprendemos la guía. Depositamos en el medio una confianza informacional y cognitiva paralela a otras confianzas prácticas sin las que seríamos incapaces de movernos. Esta descarga de tareas computacionales es continua y continuamente inadvertida, es uno de las formas más significativas en las que los humanos nos involucramos en nuestro entorno: hacemos que él nos resuelva problemas que de otra forma nuestras limitaciones mentales serían incapaces de resolver. Hacemos que nuestro entorno se pueble de andamios que sostienen nuestra fábrica humana.

El segundo paso ocurre circunstancialmente con el tiempo, la adaptación y el aprendizaje:

INTERNALIZACIÓN DE PROCESOS DE DESCARGA: Algunos de estos procesos de descarga se internalizan y convierten en módulos interactivos con las redes neuronales, cuando nuestra mente responde a ellos de manera automática, sin mediar procesos de inferencia o reconocimiento de fiabilidad.

Este segundo paso ocurre solamente con ciertos sistemas de computación externa que se convierten en auténticas prótesis mentales, más allá de su anterior estatus de herramientas cognitivas. Chalmers y Clark ponen el ejemplo de una persona, Otto, que padece Alzheimer y usa un cuaderno de notas para descargar ahí toda la memoria a corto plazo que necesita. Cuando decide ir a visitar una exposición en el MOMA, (después de consultar su cuaderno), mira allí también la dirección del museo en la calle 53. Lo que hace del cuaderno de Otto una memoria externa es que la consulta la hace sin mediar inferencia sobre su fiabilidad, como hacemos normalmente con la consulta de nuestra agenda, a la que precede en deseo de recordar, la inferencia de dónde encontrar la información, etc. Otto usa su agenda como nosotros nuestra memoria. En este sentido, el cuaderno de Otto es una prótesis, un sistema de ayuda esencial y una prolongación de una función que no podría ser realizada en otro caso. Por eso, debido al *Principio de Paridad*, lo consideramos tan componente de la mente como lo son las memorias internas. Quienes hayan visto la película *Memento* recordarán cómo el protagonista intenta preservar el sentido de

su vida apoyándose en textos tatuados en su cuerpo y unas cuantas fotografías que dan una continuidad a sus acciones que su lábil memoria de trabajo no es ya capaz de mantener. Que su vida sea lamentable no impide que le consideremos una persona que intenta por todos los medios ordenar su vida ordenando su mente y ordenando el mundo, como hacemos todos los demás. Nuestra compasión por el nace del hecho de que sabemos cuán fácilmente se manipula a una persona así, exactamente como ocurre con los niños.

Clark (2003) ha insistido en el carácter híbrido y de “ciborg” que tiene la mente humana. Nuestra mente social solamente es posible en un mundo lleno de herramientas cognitivas. La primera de ellas fue el lenguaje, que antes o después o por debajo o por encima de la estructura gramatical, es un sistema de herramientas para almacenar información, facilitar el procesamiento etc. Pero también y sobre todo para nuestra principal habilidad como animales, la de manipular la mente de los otros infectándola con nuestras ideas, deseos, creencias, expectativas y planes. Hasta qué punto se producen transformaciones de la mente como producto de los procesos de descarga/internalización es algo que podemos calibrar si pensamos en los primeros procesos del lenguaje. Vigotsky creó la idea de una zona de desarrollo próximo en la que los humanos alcanzan cotas de habilidad con ayuda de los otros que no lograrían sin ella. El lenguaje es el caso principal. El lenguaje crea un medio de interacción con otros que permite, por ejemplo, el soliloquio, las capacidades de simulación interna de la mente de los otros y con ella todos los lazos sociales. La tesis que estamos sosteniendo es independiente de cuántos sean los componentes innatos del lenguaje. Pues la inserción de sea cual sea la capacidad innata, de adecuarse a un medio externo social es un caso de hibridación computacional con el medio.

Lo que se extiende más allá del cráneo y de la frontera de la piel son los vehículos del contenido, que se convierten en prótesis computacionales. Hasta dónde llega o por donde se extiende la mente humana es algo que tenemos que examinar individuo por individuo, observando el itinerario de su historia cognitiva.

1.1. Problema 1: el contenido intrínseco

Uno de los argumentos más interesantes que se han dirigido contra la mente extendida proviene de quienes conceden casi todos los puntos pero consideran que los sistemas híbridos carecen de la principal marca de lo mental, el “contenido intrínseco”(Adams-Aizawa 2001), pero está detrás de esta objeción un espíritu de crítica a la IA en la línea de Searle y de todos los que intuyen que la particular forma biológica de nuestra especie introduce un elemento normativo. La objeción es que el punto en discusión no es si los cuadernos de notas son internos o externos, sino

si tienen o no contenido intrínseco, que solamente tienen, que sepamos, se dice, los sistemas neuronales. Pues incluso los ordenadores solamente almacenan conceptos en tanto en cuanto sean seres humanos quienes los utilicen, pues al final, solamente el contacto con neuronas hace que ciertos sistemas adquieran contenido.

Hay en esta crítica un núcleo argumentativo semántico, acerca de la naturaleza del contenido mental, y una periferia conceptual que tiene que ver con el lugar de la “folk psychology”, de la psicología natural, en nuestra idea de lo mental. El argumento técnico es el semántico, pero lo que presta fuerza a su crítica es el compromiso ontológico que está unido más o menos estrechamente a nuestra concepción intuitiva del mundo mental. El argumento semántico se dirige contra los autores que A&A llaman “transcraneales”. Estos autores serían, además de Chalmers y Clark, Dennett y Hutchins, y en general todos aquellos que defienden una idea de conocimiento distribuido en el medio (físico y social). Esta distribución, que implicaría concebir que los límites de la cognición se extienden más allá del cerebro debería cumplir según A&A la siguiente condición:

Las fronteras de la cognición deben determinarse encontrando la marca de lo cognitivo y luego viendo qué clases de procesos del mundo portan esa marca. Siguiendo este método, vemos que, como cuestión contingente de hecho, los procesos cognitivos que encontramos en el mundo real ocurren todos limitados al cerebro. Parece ser precisamente un hecho empírico el que los procesos cognitivos no sean procesos transcorpóreos (Adams-Aizawa 2001, p. 46)

La idea es que la “marca de lo cognitivo” tiene la misma extensión espacial que el alcance de la cognición. Esta marca sería la de que el estado informacional tuviese “contenido intrínseco”, a diferencia de todos aquellos estados que tienen contenido derivado y dependiente de otros sistemas que sí tienen contenido intrínseco. Así, los artefactos epistémicos como representaciones, textos, ordenadores y, en general, los artefactos que portan, producen o procesan representaciones, no tienen contenido intrínseco en sí mismos sino derivados de un cerebro que los ha diseñado. La idea de “contenido intrínseco” ha sido defendida principalmente por Dretske, por Searle y en general por quienes tienen una visión fuertemente normativa de la biología como fuente de la normatividad semántica. Adams y Aizawa no tienen inconveniente en permitir que algunos instrumentos sean extracraneales, pero no aceptan que porten por sí mismos la marca de lo mental que es el contenido intrínseco:

Las sartas de signos significan lo que significan en virtud de asociaciones convencionales entre ellas y las palabras del lenguaje. Los numerales de varias clases representan números en virtud de acuerdos y prácticas sociales; la capacidad representacional de la ortografía deriva, en este sentido, de las capacidades representacionales de los agentes

cognitivos. Por el contrario, los estados cognitivos de los agentes cognitivos normales no derivan sus significados de convenciones o prácticas sociales” (p. 48)

Adams y Aizawa comparan el proceso biológico de la digestión con la cognición extendida. Aceptan, por ejemplo, que algunos procesos digestivos puedan realizarse extra-corporalmente. Por ejemplo, aquellos animales coprófagos, como por ejemplo las gallinas, comen heces porque las bacterias presentes en ellas facilitan la ruptura de proteínas que permiten luego ser asimiladas en la digestión. Pero cuando se refieren a lo que podríamos llamar computación distribuida en el medio consideran que, a diferencia de la digestión, la cognición se parece más a la contracción muscular, que solamente puede realizarse en los músculos. Si, como ellos mismos han advertido, es una cuestión contingente y empírica el que la cognición se realice en el cráneo (consideran que es posible pero inverosímil que pueda realizarse fuera), ¿por qué esa cuestión de principio acerca de la imposibilidad de que algún contenido mental sea procesado externamente? La razón parece estar en que presuponen una dicotomía profunda e inamovible entre lo natural y lo artificial como fuente de contenido. Lo intrínseco correspondería a lo natural, mientras que lo derivado correspondería a lo artificial.

Son muchas las razones para dudar de esta dicotomía, pero quizá las que se acercan más al centro de la argumentación sobre el contenido tienen que ver con la particular evolución de los primates y en especial de los homínidos. La razón es que nuestra especie evolucionó en un medio especialmente culturizado, sin el que hubiera sido imposible la emergencia de la mente (Broncano 2004). En este medio, se produjo una coevolución de las capacidades metarrepresentacionales, que darían origen a la teoría de la mente, y de ahí a la capacidad para crear convenciones, de las que deriva el propio lenguaje, y la capacidad para preservar medios culturales favorables a la sintonía con los sistemas simbólicos. Sin una presencia humana fuerte en los procesos de maduración cerebral no existe un entorno próximo suficientemente rico como para la adquisición del significado, independientemente de las capacidades innatas para adquirir gramáticas. Lo mismo ocurre con los medios técnicos: la capacidad para crear planes y los artefactos con estructura compleja, como son ya los útiles líticos, coevolucionan en interacción. En cierta forma, pues, nuestra mente es ya un producto artificial. No hay un momento originario en el que se establezca una frontera entre lo intrínseco y lo derivado.

Si esta distinción se pone en cuestión en lo que respecta al origen normativo del significado, de acuerdo a los promotores de la idea de un momento biológico normativo originario, la dicotomía entre lo natural y artificial se relativiza mucho y se convierte en algo contextual y dependiente de consideraciones pragmáticas y convencionales. De hecho no es tanto la dicotomía entre lo natural y lo artificial lo que está involucrado en qué parte de la mente puede extenderse al medio más allá de la

frontera epidérmica, sino qué partes de la mente son personales y cuáles subpersonales. Dennett (1991) distinguió muy inteligentemente entre los contenidos y cómo son producidos por el cerebro. Así como ya sabemos que nuestras imágenes tal como se producen en la percepción no son “imágenes en la cabeza”

Deberíamos establecer la misma distinción para el tiempo: el momento en que se produce una experiencia en el cerebro debe distinguirse del momento en que ésta parece producirse (...) La representación del espacio en el cerebro no siempre hace uso de el-espacio-en-el-cerebro para representar el espacio y la representación del tiempo en el cerebro no siempre hace uso de el-tiempo-en-el-cerebro” (Dennett 1991, p. 145)

La distinción es entre los vehículos del contenido y los contenidos mentales. Los contenidos aparecen en el nivel personal, en el que se produce la comprensión y la semántica, no en el nivel subpersonal, que procesa información, y no hay por qué postular que las propiedades de los contenidos sean una representación isomórfica de los vehículos y viceversa (Hurley 1998). El punto es entonces si pueden o no existir vehículos que se extiendan más allá de la piel y que sin embargo sean portadores de contenido. Pues es en la respuesta a esta pregunta en donde se sitúa el centro de la cuestión de la mente extendida. Andy Clark, en al menos dos escritos (Clark 2005 y 2006) refuta la cuestión de principio implicada en la crítica acudiendo a esta distinción. Su tesis no es que todos los vehículos sean externos al cerebro, sino que pueden darse procesos de asimilación de vehículos externos que por ello se convierten de esta forma en verdaderas prótesis mentales y no sólo corporales, de modo que extienden las funciones mentales más allá de la frontera de la piel de tal manera que ya no pueden ser concebidas solamente como herramientas. No al menos en el sentido en que prótesis como las gafas o las fundas dentales no pueden ser concebidas como herramientas sino como compleciones de funciones corporales realizadas externamente.

El proceso de externalización de la mente se da, de esta forma en los dos pasos diferenciados de descarga e interiorización. La descarga se realiza en una parte del medio adecuada de forma robusta para la tarea: por ejemplo, un aparato de computación que es usado permanentemente, una libreta, etc., es decir, un sistema que pueda acoplarse a la computación. La progresiva interiorización del artefacto pone en cuestión el que las barreras del cráneo o de la piel sean tan adecuadas como parece para definir la frontera entre lo mental y lo no mental. La interiorización no es en absoluto un proceso ajeno al cerebro, es de hecho una forma en la que su constitución plástica se adapta al medio de mismo modo que “modulariza” otras muchas tareas aprendidas constituyéndolas como automatismos. Se produce así una inestabilidad respecto a qué es lo mental que afecta no tanto al dentro/fuera sino a los aspectos personales/subpersonales de la vida mental, pues, claramente, los procesos que se descargan en el medio pertenecen a los procesos computacionales que son no conscientes aunque causalmente activos en la mente.

1.2. Problema 2: la autoridad epistémica

En segundo lugar vamos a considerar una nueva clase de dificultades que ha sido desarrollada por Jesús Vega (2006) y que no afectan, como en el caso de Adams y Aizawa a la naturaleza del contenido sino a la autoridad epistémica en los procesos mentales. La tesis de Vega se resumiría en que todo lo que podemos conceder a la tesis de la mente extendida es que se trata de un control sobre artefactos cognitivos o artefactos epistémicos en el que “el modelo de mente extendida y distribuida instaure un modo en que la autoridad procede directamente a través de la deferencia en otros (recursos humanos o artificiales)” (p. 31). Según Vega habría una diferencia radical en el modo en el que adscribimos autoridad epistémica a los contenidos en primera persona del modo en el que adscribimos autoridad a los contenidos provenientes de otras personas o de artefactos epistémicos construidos por otros, en donde la autoridad no se basa en la familiaridad sino en la confianza en otros o en la fiabilidad del instrumento. Una mente, desde el punto de vista normativo, sostiene Vega, consiste esencialmente en una capacidad subjetiva para reconocerse a sí misma como fuente de autoridad epistémica:

La mente *subjetiva* se reconoce como fuente de autoridad en la medida en que es capaz de evaluar racionalmente las jugadas que hace de aceptación o de rechazo *responsable*. Uno no tiene por qué comprometerse con una imagen del espacio de lo mental según la cual somos infalibles e incorregibles epistémicamente, ni admitir un modelo de acceso observacional (de tipo introspectivo) a los estados mentales, ni hacer de la conciencia la propiedad especial de todo estado mental que explica el acceso y su autoridad para reconocer cómo las peculiaridades del acceso mental están estrechamente vinculadas con el tipo de autoridad epistémica que asociamos con la idea de subjetividad (Vega 2006, p.5).

El argumento tiene un reconocible aroma kantiano. Lo que nos convierte en sujetos es la autonomía, la capacidad de dotarnos a nosotros mismos de reglas y reconocer en esas reglas su fuerza normativa. Y ese reconocimiento, que conlleva la asunción de responsabilidad, solamente puede hacerlo una mente dotada de subjetividad y conciencia.

Concede Vega que la tesis de la mente extendida sería una tesis que es aceptable nada más que como una curiosidad dentro del marco de las nuevas aproximaciones a lo mental que son conocidas como “mente incrustada y encarnada” (embedded and embodied mind), o mente situada en el medio y corporeizada o constreñida por los esquemas sensorio motores corporales. Estas aproximaciones contemporáneas, que fueron anticipadas por Heidegger y Merleau-Ponty, que sólo recientemente han sido incorporadas a las ciencias cognitivas gracias, entre otras, cosas a las nuevas perspectivas que abren nuevos conocimientos sobre las redes

neuronales y su forma especial de procesamiento de la información, como el conexionismo y la aplicación de modelos dinámicos de programación o los algoritmos genéticos que tienen cierta forma de autoprogramación, o las perspectivas de la nueva robótica que se apoya en la idea de aprovechar el medio como fuente de computación. Son aproximaciones que se enfrentan a muchas de las características de la visión cartesiana de una mente esencialmente “privada” interna. Vega no rechaza estas aproximaciones, pero sostiene que la idea de la mente extendida no va más allá de lo que va el situacionismo y el “encarnacionismo”:

Si la situacionalidad insiste en el papel del entorno como esencial en la naturaleza de los estados y procesos cognitivos, en la medida en que este entorno incluye estas extensiones, podría afirmarse que la cognición extendida no es sino uno de los modos paradigmáticos de la cognición situada (Vega 2006, p. 17).

Jesús Vega reconoce que la especie humana es capaz de construir un medio artificial lleno de artefactos epistémicos que se definen funcionalmente por su capacidad de éxito como portadores o procesadores de información. Sin embargo, por muy perfectos que sean, los artefactos cognitivos son radicalmente dependientes del sujeto:

Los artefactos cognitivos cumplen exitosamente su función en la medida en que los agentes comparten una serie de competencias y habilidades que conducen no sólo a interpretar correctamente la información que hacen disponible sino también a coordinarse adecuadamente con ellos. Estas capacidades son de distinto tipo, pero hay tres elementos ineludibles: en primer lugar, requiere una competencia semántica de interpretación; en segundo lugar, está involucrada una comprensión de la relación representacional mediante el dominio de capacidades metarrepresentacionales; en tercer lugar, se requiere una capacidad evaluativa del modo en que se adquiere y se modifica *fiablemente* la información. Todas ellas son requisitos del uso eficiente de las herramientas cognitivas. Esto quiere decir que el *acoplamiento* con las extensiones de la mente depende esencialmente de cómo se gestionen estas capacidades (Vega 2006, p. 25).

Ahora bien, estas capacidades que debe tener el sujeto para que en interacción con los artefactos alcancen éstos a cumplir su función tienen un origen que no puede ser externo y situado:

...su uso depende de la presencia de un tipo especial de recursos internos, de representaciones “desacopladas” y de propósito múltiple, no dependientes de la situación y del contexto, sino aprendidas y gestionadas independientemente de las condiciones concretas del entorno. Sin ellas, sería imposible dar cuenta del éxito en el uso de los artefactos epistémicos (Vega 2006, p. 25).

El punto que divide la idea de lo mental no es tanto el éxito funcional de los aparatos cuanto la capacidad normativa que, según Vega, solamente puede dar lo mental: “Realmente, lo que está en juego es menos la fiabilidad, el tipo o la calidad en el acceso cuanto las características de las *evaluaciones* que son ineludibles para ajustar el éxito de los artefactos cognitivos” (26). Hay tres características que, según Vega, impiden que consideremos las prótesis cognitivas como parte de la mente

1) Las evaluaciones normativas exigen capacidad metaintencional y metarrepresentacional:

El tratamiento metaintencional es genuinamente *metarrepresentacional*: no basta con que el elemento externo haya sido producido intencionalmente; debe situarse en un contexto interpretativo especial en el que el mismo elemento material soporte el hecho de que en un mundo notional del agente productor de la huella se sostiene otro hecho (Vega 2006, p. 26)

Es decir, el sujeto debe hacerse presente en la capacidad interpretativa de un artefacto cognitivo en tanto que artefacto cognitivo. Esta presencia se da en forma de la “*comprensión* que el agente manifiesta a través de su coordinación sensorio-motriz y su manipulación. Sobre esta comprensión metarrepresentacional descansa la explicación del éxito funcional de las representaciones externas. Es más, aunque el tipo de tareas cognitivas y la realización funcional de las mismas sólo pueda ser explicable por medio de los recursos externos con los que el agente se *integra*, las atribuciones de *mentalidad* y de *intencionalidad* están soportadas por el agente mismo, pues de él depende el control metaintencional de los procesos” (27).

2) La evaluación epistémica se hace con respecto a estados o habilidades propias, y no con respecto a los procesos causales del aparato. Así, Vega pone el ejemplo del reconocimiento de un error en el funcionamiento de un aparato.

El agente tiene que establecer las condiciones bajo las cuales podría poner en duda sus compromisos anteriores con los resultados de los cálculos. Y posteriormente debería poder descubrir que el error procede de un defecto físico en el aparato. El defecto físico *afecta* al resultado de sus aceptaciones o rechazos, y el agente explicaría sus errores (ante otros) como causados por tal defecto. La red normativa de estados cognitivos (intencionales) no incluye, sin embargo, una consideración también normativa del aparato (Vega 2006, p. 28).

Para que el agente reconozca autoridad epistémica, por ejemplo, a lo escrito en un cuaderno de notas, como es el caso de Otto que padece pérdida de memoria a corto plazo:

Necesita tomar lo allí escrito como *sus propios estados mentales* y no como una parte de sí que escapa completamente a su control. ¿No se admite con ello que la cognición se ha distribuido y que el cuaderno y lo escrito en él forman parte de su mente? Sin duda, hay un aspecto distribuido en el modo en que la función de memoria es ejercida por Otto, pero su valor cognitivo depende en último término de cómo Otto se trate a sí mismo como estando dispuesto a aceptar los resultados. En último término, ahí se cierra el círculo de la intencionalidad, lo que integra sus estados mentales y lo que permite gestionar adecuadamente su control. Sigue siendo el depositario de la autoridad *cognitiva y epistémica* (Vega 2006, p. 28).

3) La fundamentación normativa exige además un momento de apercepción kantiana:

... la idea de una cierta unidad y coherencia del espacio normativo de lo mental no se agota en la posibilidad de supervisión en forma de estados de segundo orden. La autoridad cognitiva y epistémica en que consiste el hacerse cargo de los propios compromisos y habilitaciones no se reduce a adoptar una actitud de segundo orden sobre un estado intencional de primer orden. Supone un reconocimiento genuino de que *soy yo* quien se hace cargo de las evaluaciones³⁹, se hace cargo también de los estados de segundo orden (Vega 2006, p. 30).

En resumen, pues, la mente extendida, si es extendida pierde lo que tiene de genuino la mente que consiste en el reconocimiento de la autoridad epistémica que conlleva un reconocimiento de propiedad de los estados mentales como propios y de los que la mente se hace responsable.

Mi impresión es que la crítica de Vega no se aleja sustancialmente del punto de A&A: la marca de lo mental no estaría ahora en la idea de contenido no derivado sino en la de autoridad epistémica. El punto en discusión sin embargo en la mente extendida no es ése, sino cuál es el locus de la subjetividad y cuál es el estatus de los estados mentales que procesan información. Vega se apunta a dos dicotomías que son básicamente cartesianas: la dicotomía entre la receptividad/espontaneidad del pensamiento y la dicotomía funcionalismo/intencionalidad: la mente es intencional en esencia aunque se base en mecanismo de éxito funcional. El locus de la mente es un espacio “interior” de razones. La tesis de la mente extendida, es cierto, implica un compromiso con ciertas formas de concebir lo mental que ponen en cuestión la primera dicotomía: ¿está tan claro que toda la parte receptiva sea funcional y sólo seamos humanos en una espontaneidad posterior en el que se reconoce la autoridad de un estado?, ¿no es más bien de un bucle continuo de interdependencias activo-reactivas en el que emerge la autoridad epistémica como haciéndose presente un principio de autoridad?, ¿no es una idea de reconocimiento demasiado dependiente de la imagen *folk*? Vega, me parece, está confundiendo las propiedades del contenido, o la descripción autofenomenológica, con las propiedades de los

vehículos, que exigen un cierto componente de heterofenomenología. Del mismo modo que las propiedades espacio-temporales de los estados mentales no son las propiedades espacio-temporales de los vehículos de pensamiento, pues el cerebro tiene sus propias formas de procesamiento.

Los vehículos no tienen por qué tener propiedades como son la meta-representación o el reconocimiento aperceptivo, y sin embargo son vehículos de la mente y de sus contenidos. Aplíquese el mismo tratamiento de Vega, por ejemplo, a las emociones y obtendríamos como resultado que las emociones, que son un sistema parcialmente autónomo de procesamiento cognitivo-motivacional³, no serían estados mentales. Lo que llevaría la tesis de Vega en último extremo es a considerar mental solamente los contenidos de la conciencia una vez sometidos a evaluación epistémica. Eso, creo, es la esencia del cartesianismo. Podríamos admitir, y de hecho admitimos, los criterios de lo mental normativos que expresan las críticas de A&A y Vega y sin embargo creer que no está afectado el núcleo de la idea de mente extendida.

Sin embargo, nos falta el por dar cuenta de lo que sí es un elemento que subyace a estas críticas y que es donde se asienta el núcleo de nuestra consideración de lo mental. ¿Cuál es el locus de la autonomía? ¿Qué nos convierte en sujetos? ¿Qué es la subjetividad? ¿Dónde existen las personas?

2. La integración de la mente

Si no es el contenido ni la autoridad epistémica lo que amenaza a la idea de mente extendida, sí hay, sin embargo, un tercer problema que me parece particularmente relevante. Está en cierto modo supuesto en los dos anteriores pero no se reduce a ellos. Se trata del viejo tema kantiano de la unidad de la conciencia traducido al más amplio territorio de la unidad de la mente. El problema de la unidad aparece en el momento de la comprensión, en el momento del juicio evaluativo (de ahí que esté implicado en las dos líneas de crítica anteriores), y también y sobre todo en la determinación agente por un sujeto que toma una decisión de acción (lo que no equivale exactamente a un juicio práctico); está implicado también en el mecanismo de la preservación de la identidad personal, mediante el cual el yo se reconoce como la misma persona que se ha embarcado en un conjunto de proyectos racionales. La cuestión que surge es si la idea de mente extendida, al ampliar más allá de

³ El sistema emocional es un sistema parcialmente autónomo que se coordina con el cognitivo, pero que en sus estadios primigenios pertenece a formas de cerebro mucho más primitivas, que evolucionaron para resolver problemas de tratamiento informacional del medio muy rápidas y robustas, pero que no alcanzan los niveles de procesamiento a través de conceptos que parece ser la marca específicamente humana.

la piel el conjunto de vehículos que procesan contenidos, no estará amenazando la unidad del sujeto y llevando a algo así como una existencia rizomática, tan atractiva a las filosofías posmodernas.

El argumento que vamos a desarrollar aquí no es inconsistente con la tesis kantiana de la necesidad de una unidad de sujeto como fundamento de lo normativo. Sin embargo, y esta es nuestra posición, podemos distinguir el problema de la unidad de la mente, el problema de la integración tal como lo vamos a denominar nosotros, del problema normativo de la unidad del sujeto y el problema normativo de la identidad personal. La neutralidad kantiana de esta división la debemos al propio Kant, que distinguió entre el problema empírico de la unidad de la mente y el problema trascendental de la apercepción (“la conciencia del yo en la percepción interna es meramente empírica, siempre mudable, sin poder suministrar un yo fijo y permanente en medio de esa corriente de fenómenos internos. Dicha conciencia suele llamarse *sentido interno* o *apercepción empírica*” (A 107)). Cuando nos referimos al sujeto y a la identidad personal estamos considerando un nivel normativo o trascendental, para usar la terminología tradicional, que no es necesariamente el de la mente “empírica” o, traduciéndolo ahora a nuestra terminología, el de la psicología sino el nivel normativo de la epistemología, la semántica o la pragmática, niveles en los que un objeto dotado de mente se reconoce a sí mismo como un sujeto y es reconocido por otros como una persona.

2.1. Problema de la individuación de lo mental

La idea de mente extensa es una refutación de la ontología cartesiana. Está formada por todos los procesos de información que contribuyen a individuar una mente. Puesto que una mente se realiza o sobreviene a un conjunto de sistemas que básicamente se definen funcionalmente como sistemas que procesan información, el conjunto necesario para la preservación de una mente individual sería parte de dicha mente, independientemente de si tales sistemas se sitúan dentro o fuera de lo que consideramos prototípicamente como cuerpo humano. Podemos distinguir el problema de la arquitectura, que es el de cómo se articulan los diversos componentes de una mente para generar lo que ésta consigue hacer, que se trata esencialmente de un problema técnico y científico, y el problema de la integración de sistemas heterogéneos, que puede ser tratado filosóficamente como un problema de identidad, mucho más abstracto que el de la arquitectura, aunque no necesariamente normativo en el sentido kantiano al que acabamos de aludir.

La cuestión que se nos plantea ahora es el de cuál es el mecanismo de individuación de una mente: la nación china de Block, la habitación china de Searle, etc... son experimentos mentales dirigidos contra la concepción funcionalista de lo men-

tal en los que se plantea ya un problema de individuación de mentes: ¿hay en la habitación china una o dos mentes? ¿Hay en la nación china una o mil millones de mentes? Me parece que estas preguntas son previas a la intuición que nos “bombea” el experimento, y que de contestar en uno u otro sentido cambiaría radicalmente la valoración de estos experimentos mentales con respecto a características particulares de una mente como son los *qualia* o la conciencia. Pues antes deberíamos tener una cierta noción de si pertenecen como propiedades a la mente en cuanto tal o simplemente a algunos de sus subsistemas.

La intuición en estos dos experimentos es contradictoria. Es como si se produjera un choque de intuiciones en el mismo guión, de ahí que surja este malestar o sensación de extrañeza al oír el experimento. ¿Qué individúa una mente más allá de una colección de sistemas de procesamiento que se articulan para cumplir ciertas funciones? De una parte está el problema de la extensión (de piel adentro, o extensa), de otra parte está el problema de qué es lo que organiza la mente como una unidad ontológica. La respuesta varía si nos dirigimos hacia un terreno puramente funcional o, si siguiendo la división anterior, creemos que en la respuesta interviene un componente normativo que trasciende a la descripción sistémica de lo mental. La respuesta funcional propondría una estrategia de determinar cuál es el criterio para la unidad de un sistema de sistemas como la mente. Incluso, cuando decimos que la conciencia, o la presencia de la conciencia, es una condición de unidad, no estamos aún dando una respuesta precisa a la dicotomía anterior, a menos que especifiquemos si lo que define la conciencia es algún elemento funcional, si el aparecer algo como experiencia es una condición suficiente o simplemente necesaria, etc. Quienes tiendan a pensar que la respuesta es normativa encontrarán que no es el funcionamiento de un sistema, por ejemplo la conciencia, el que determina la mis-midad, sino que las condiciones hay que buscarlas en cierta forma de comportarse el sistema entero tales que se produce un autonomía real del sistema. Otra cuestión es respecto a qué consideramos que el sistema debe ser autónomo. La respuesta, en la línea kantiana, es la de que su conducta esté determinada por los propios juicios. O podemos definirlo en término espinosianos como “causa sui”.

Concedemos el estatus de persona a seres que aún no han desarrollado sus capacidades mentales como es el caso de los niños. Concedemos también el estatus jurídico, moral y afectivo de personas a seres que han perdido sus facultades mentales, como ocurre con quienes padecen un coma irreversible u otras graves disminuciones de la capacidad mental. No le concederíamos, sin embargo, este estatus si no supiéramos que en el futuro podrán llegar a tener mente, o que la han tenido, o que tendrían que haberla tenido de no mediar estas perturbaciones. Por otra parte, sin llegar a los extremos que presentan estos casos, nos encontramos ante mentes que ejercen sus funciones parcialmente, o que lo hacen de forma disfuncional sin que por ello afecte en absoluto a su estatuto de personas. En consecuencia, la mente y

las personas se individúan de formas distintas, aunque sean interdependientes en algunos aspectos. Puede que compartan bases de superveniencia o puede que las relaciones de dependencia sean mutuas y se produzcan de forma muy compleja (sometidas a condiciones límite y a restricciones en un orden de heterogeneidad ontológica).

Mi posición, como ya he insinuado anteriormente, es distinguir claramente los dos niveles: la unidad de la mente es una cuestión distinta de la identidad personal. Mientras que el problema de la unidad de la mente es un problema de integración funcional de sistemas, y el criterio de individuación debemos encontrarlo en los criterios de suficiencia de la integración, el problema normativo de qué convierte a un objeto con mente en un sujeto es un problema que trasciende lo funcional y nos lleva a otros territorios. Así, mientras que la noción de lo mental puede ser tratada en términos de sistemas, las ideas de subjetividad, de sujeto y de persona exigen la alusión a espacios sociales de compromiso y responsabilidad sin los cuales no tienen sentido. Dicho de otra forma. Mientras que la mente admite, y quizá exige, un tratamiento hetero-fenomenológico, o en tercera persona, independientemente de que sea necesario acudir en ciertos momentos al nivel auto-fenomenológico, en primera persona, en los informes introspectivos, en el tratamiento del sujeto, persona y subjetividad la autoridad de la primera persona debe ser preservada, por más que necesitemos el nivel interpersonal y la tercera persona para dar cuenta de las normas, espacios de razones, etc.

2.2. La integración de funciones: dos aproximaciones

Una mente es, para decirlo rápidamente, un sistema de sistemas de control, o una red de sistemas de control que soporta prácticas y conductas que calificamos como inteligentes. Tendríamos que distinguir, para seguir coherentemente nuestro camino, entre conductas inteligentes y conductas racionales. Mientras que la inteligencia hace alusión a la capacidad de solución de problemas en términos de éxito funcional, no todas las soluciones de problemas son racionales. Cabe pensar, al menos por mor del argumento, que conceptualmente son distintas propiedades. Podríamos pensar en seres inteligentes e irracionales. Otra cuestión distinta es que nos planteemos la pregunta de qué clase de inteligencias pueden soportar racionalidad. Pero de nuevo nos encontramos con la distinción anterior entre lo normativo y lo descriptivo. Así, es posible que necesitemos inteligencias que sean capaces de reconocer la mente de otros, de reconocerse a sí mismas, etc., es decir, de resolver problemas que pueden plantearse en términos normativos. Pero ello no obsta para que la inteligencia pueda tener sus propias vías ligadas a la naturaleza de la mente. Por supuesto se me preguntará con intención crítica si por inteligencia estoy entendiendo el pasar el “Test de Turing” y con él la capacidad conversacional.

Ciertamente, en el Test de Turing está implicado un elemento deíctico que ancla la inteligencia a cómo la reconocemos los humanos dotados ya de todos los elementos normativos que nos sitúan en el espacio de las razones. En este sentido, la separación ficticia de inteligencia y racionalidad no permitiría a esos seres pasar un Test de Turing en el que los jurados fuesen humanos, en el sentido completo que lo somos los sujetos. Pero, también por mor del argumento piénsese en los niños, en sujetos con déficits metarrepresentacionales, etc. a quienes en principio les consideramos dotados de una mente sin considerarles seres racionales, aunque desde luego mi pretensión de separar lo psicológico de lo normativo no se refiere a estos seres como prototipos, sino más bien a inteligencias concebidas abstractamente. Por ejemplo, inteligencias artificiales que evolucionasen fuera del espacio humano intersubjetivo.

Ahora bien, planteada la cuestión de la identidad de una mente en estos términos podemos ya observar que las condiciones de individuación no tiene por qué ser las mismas que el cuerpo: dependen de la estructura de los bucles de realimentación de los sistemas de control. De hecho, la identidad funcional, ya en el nivel fisiológico, no tiene por qué confundirse con la identidad corporal. Antes aludimos a animales que completan su digestión ingiriendo heces que son alimentos porque previamente han sido transformados por las bacterias del medio. Cabe pensar en animales absolutamente coprófagos para los que su dependencia del medio fuera absoluta en lo que respecta a la fisiología de la digestión, por lo que no podríamos decir que en esa especie la digestión es el proceso que realiza el sistema digestivo, entendiendo por tal el puro sistema de órganos orientados a la digestión. La integración funcional digestiva supone una integración funcional de los órganos que incluye los mecanismos funcionales fuera del cuerpo, aunque, claro, también y de forma dominante, todos los mecanismos funcionales intracorpóreos. Aquí el punto está en la integración, que podemos considerarla como un proceso o dinámica del sistema completo, que entraña procesos de coordinación y ajuste funcional bajo ciertas condiciones de éxito biológico que están definidas también en el diseño funcional del sistema.

En este sentido, hay una ambigüedad en la dicotomía sistema/planta de la que hablan los ingenieros de control o de la dicotomía medio/mente que se presupone en la mayoría de los estudios psicológicos. Podemos considerar como pertenecientes al sistema de control a todas las partes que nos permiten definirlo como sistema: la fuente de datos, el motor de inferencias (o las múltiples fuentes de datos y motores de inferencias, el almacén donde se hallan los objetivos y planes del sistema, los detectores de información, etc. La distribución particular en el espacio de un sistema de control puede ser muy extensa en un sentido nada trivial. O, por el contrario, podemos pensar que el cuerpo originario tiene algo más que una función de soporte, sino que, de alguna forma, contribuye de forma esencial a la individuación del

sistema, de manera que si fuésemos algo así como robots, el cuerpo tal como nos ha sido dado por la evolución contribuye de manera determinante a la individuación de la mente. Pero el ejemplo anterior del cuerpo nos muestra que la individuación del sistema es esencialmente funcional y podría no coincidir con el diseño biológico o puramente mecánico, que es el vehículo de preservación del sistema, sino con el diseño funcional, con el fenotipo extendido, que solamente existe si en el medio se produce una cooperación para definir el sistema funcionalmente. Si en el medio al que se refiere nuestro ejemplo, el de los animales exclusivamente coprófagos, tales seres no encuentran heces, independientemente de que no logren sobrevivir, sobre todo no logran completar su sistema digestivo.

Una fábrica llena de sistemas de control no es una mente en el sentido integrado, por más que su comportamiento sea extremadamente complejo y por más que interactúe con el medio de forma inteligente (al menos en apariencia). ¿Cuándo consideramos el sistema como un sistema unificado en relación a lo que hemos llamado inteligencia en el sentido psicológico del término? En realidad nuestra separación ficticia de inteligencia y normatividad racional solamente tenía como objetivo abrir la posibilidad de un continuo de sistemas mentales o sistemas de funciones integradas inteligentes que se jerarquizan respecto a las tareas que son capaces de desarrollar. Es menos importante el saber si alcanzan la categoría de mentes humanas que saber si alcanzan la categoría de tener una mente unificada en un sentido de dinámica funcional integrada.

La noción de integración de sistemas de control nos permite definir un criterio de unidad mental de forma muy abstracta, sin considerar aún si lo que nos produce es una o varias mentes humanas. La integración puede darse de múltiples formas, pero la que a nosotros nos interesa es la que nos permite diferenciar de alguna forma el dentro/fuera del cuerpo. Humphrey (2000) toma como criterio de conciencia el poder tener sensaciones. Este criterio implica, según Humphrey distinguir entre percepciones, mediante las que el ser adquiere una información sobre el medio externo, y sensaciones que son actividades (no objetos) mediante las que el ser obtiene información sobre sí mismo, sobre algo que le está ocurriendo a él. Las sensaciones tendrían ciertas propiedades como la *auto-pertenencia*, la *localización corporal* (como deíctico referido a una parte del cuerpo), la *modalidad cualitativa* que identificamos con el qualia de esa sensación, la *presencialidad* o aparición en tiempo presente y la *inmediatez* fenoménica. No vamos a discutir estas propiedades una por una. El punto central es que son propiedades de una actividad que implica una cierta forma de autorreferencia del sistema que podría en principio ser realizada por un sistema artificial que hubiese sido diseñado para tener un cierto control de sus acciones dividiendo la información que recibe del mundo de la que recibe sobre sí mismo. Si tal ser es posible no es una cuestión que podamos discutir aquí⁴. Lo que

⁴ Dennett (2000) afirma que sí a condición que tenga cierta forma de diseño *ad hoc*.

importa es que la cuestión de la mente la llevamos a una cierta forma de integración sistémica que entraña una distinción entre percepciones y sensaciones. ¿Podría este criterio ser compatible con la idea de mente extendida? No veo la razón de por qué no. En el caso de las mentes protésicas, probablemente tendríamos alguna forma particular de modalidad sensoria extraña, pero no entiendo por qué un sistema integrado no puede distinguir el dentro (me ocurre a mí) /fuera (percibo algo que ocurre en el medio), claro, en el caso de que haya integrado una prótesis de computación externa. Pensemos en Otto, el paciente de Alzheimer y su cuaderno de notas. Imaginemos que el cuaderno sufre cambios que afectan directamente a la capacidad de memoria de Otto: desde luego Otto no puede experimentarlos en forma de dolor u otras formas de sensación cuyas funciones sean la detección de cambios en los tejidos corporales. Pero puede detectar: me está ocurriendo algo en la memoria en un sentido estricto y funcional. Si eso ocurre, tenemos una mente extensa unificada e integrada.

Tenemos pues que la solución al problema de la individuación de la mente es la búsqueda de una solución al problema de la integración jerárquica de los sistemas más que una solución al problema de la “identidad” mental. Una de las cuestiones que suscita la psicología es precisamente la de la integración de los sistemas en condiciones de operación no normales, por ejemplo en los casos de disociación de personalidad, de interferencia de unos subsistemas con otros, etc. En estos casos, la amenaza es precisamente la de la fractura de la voluntad, la incoherencia, etc. Hay numerosas amenazas a la integridad de la mente: las incongruencias más interesantes nacen cuando se presentan tensiones entre sistemas funcionales completos que, sin embargo, tendrían que actuar armónicamente, como ocurre, por ejemplo, entre el sistema emotivo y el sistema cognitivo, entre el sistema visual y el auditivo, etc. A veces, sistemas modulares completos entran en competencia unos con otros. El punto central es si existe o no un criterio por el que podamos hablar de integración funcional. Quizá el criterio de Humphrey es suficiente, no lo sé: es una propuesta que se mueve en la línea de una forma de naturalización de lo mental que no implica un *gap* o hueco entre lo funcional y algo extraño que sería lo cualitativo o fenoménico intrínsecamente⁵. Pues el lugar de la discusión es qué actividades de con-

⁵ De hecho tiene problemas en lo que respecta a las propiedades fenoménicas y su relación con las propiedades científicas, meramente informacionales, de la conciencia en la sensación y percepción (véase Eilan 2000), pero esas dificultades, que tienen que ver con la tensión entre la imagen científica y cotidiana o fenoménica de la conciencia, no afectan al núcleo de su criterio que puede concebirse como la emergencia de una conciencia de “separación” del dentro/fuera, sin que tenga por qué ser pensado en términos puramente espaciales, podría serlo, como de hecho ocurre en los mamíferos, como una conciencia mío/no mío (véase Damasio 2000). El punto que amenazaría nuestro argumento sería el que en la conciencia fenoménica ya estuviese implicado un componente fuertemente normativo como el que nos exigen los criterios epistemológicos. Pero nuestra posición es que nos movemos aún en un territorio empírico en el que la conciencia primitiva de separación corporal puede ser adscrita a animales que no tienen la sensibilidad hacia lo normativo que tienen los humanos.

trol son distinguidas como control de la propia mente, de los propios límites, y cuáles son actividades de control del medio externo. Esta conciencia de separabilidad y pertenencia que tienen los accesos informacionales al mundo puede coincidir con el diseño corporal, como ocurre normalmente, suponiendo que en el cerebro hay representaciones y mapas corporales que permiten una integración de funciones en esta dirección, o quizá, como postula la idea de mente extensa, dada la enorme plasticidad de lo mental, cabe una concepción “abierta”⁶ de la extensión de la mente, en la que los límites no son “naturales” sino funcionales. La cuestión central sigue siendo, pues, la de la integración cognitiva de los sistemas de sistemas de control que constituyen la mente.

3. La subjetividad como resultado

El siguiente punto es entonces el de encontrar el *locus* de propiedades más complejas de orden normativo como son las que asociamos a la subjetividad, a la intencionalidad y a la agencia. La aparición de la conciencia y el dominio de la subjetividad en una escala evolutiva presenta numerosos problemas de hipótesis funcional que nos ayudarían a comprender las principales fuerzas que impulsaron la trayectoria del reforzamiento de la conciencia.

La distinción apariencia/realidad y la emergencia de una mente representacional.

La aparición de un espacio subjetivo y la construcción compleja de qualia como identificación de los rasgos fenomenológicos del mundo (construcción de una fenomenología como hetero-fenomenología).⁷

La representación idealizada de las trayectorias narrativas, imputando al yo la responsabilidad de esa trayectoria.

Estas tres características están implicadas en la formación de la mente racional que responde no solo a los retos informacionales y funcionales del medio sino a las propiedades normativas que emergen en su contexto (externo o interno). Pues parece claro que nuestro modo de involucrarnos en planes racionales y trayectorias vitales exige una mente representacional, una mente sensible a propiedades fenoménicas, que son las que componen en buena medida nuestra noción de realidad aparente y, por último, aunque no menos importante, nuestra capacidad para encontrar sentidos a trayectorias largas de la existencia, de las que está hecha la identidad perso-

⁶ Varias conversaciones sobre el tema con Jesús Vega me han llevado a pensar que la idea de “mente abierta” puede inducir a menos errores que la de mente extendida. Pero tal vez sea solamente una cuestión nominal.

⁷ En este espacio incluimos también el “sentido” del yo que aparece en la agencia y en las formas primitivas de autorreferencia.

nal. En este nivel aparecen todas las exigencias normativas que hemos comentado ya en la respuesta a los críticos de la mente extendida. Aquí es donde puede aparecer un replanteamiento del problema de la mente extendida. Pues podría postularse que aunque funcionalmente estuviese integrado un sistema que hubiese asimilado subsistemas de procesamiento pertenecientes al medio, en nuestra concepción normativa de nosotros mismos, la conciencia dentro/fuera, como conciencia normativa de nuestro propio cuerpo es una condición a priori de nuestra sensibilidad, y, por consiguiente, la mente extendida, de serlo, aparecería solamente en los niveles puramente primarios de la constitución mental, pero no en los niveles normativos en los que se constituye una mente racional y en donde se constituye la agencia. Me parece que esta objeción está en la base de las anteriores, una vez que despejamos la posible confusión que hubiese entre adscribir propiedades al contenido y adscribirlas a los vehículos de dicho contenido.

La discusión se plantea, me parece, en el punto en el que convergen una cierta idea de lo mental que está presente en nuestro lenguaje cotidiano y el componente normativo que resulta de un examen filosófico, que tiene que tener en cuenta, para seguir la vieja idea de Sellars, tanto la imagen científica como la imagen manifiesta y, si es posible, tejerlas en una mismo tejido conceptual.

En primer lugar, debemos separar los compromisos ontológicos que están anclados con más o menos profundidad en la imagen manifiesta de la idea normativa, que bien podríamos comprenderla en términos kantianos como condiciones de posibilidad de nuestra constitución como seres dotados de una cierta forma de mente. Pues bien, los compromisos ontológicos de la imagen manifiesta, de nuestra “folkpsychology” son ambiguos respecto a la fuerza normativa. Algunos elementos parecen bien justificados, otros mucho menos. Por ejemplo, parece que las nociones de creencia, deseos, intenciones, etc. que parecen formar parte de nuestra concepción de lo mental no pueden ser revisadas sin revisar completamente la propia epistemología que subyace a la ciencia, puesto que estaríamos tocando los elementos más sensibles de nuestra idea de racionalidad. Emerge así una cierta idea de lo mental que ha sido conceptualizada por los teóricos como “Teoría de la Mente”, como si fuera un componente teórico implicado en nuestras prácticas lingüísticas. Pero, solamente por citar un debate bastante bien conocido, esta concepción no es capaz de decidir si es una “teoría” o esquema conceptual que aplicamos a los datos, en este caso, a los datos conductuales de los otros para explicar o predecir su conducta, o, por el contrario, es una capacidad interna de simulación de la mente ajena. El debate “Teoría vs. Simulación” no es decidible dentro de la imagen manifiesta, y sin embargo es central en sus compromisos ontológicos. Otros elementos como qué es un sujeto o qué es una persona, sin embargo, sí lo son o es quizá el análisis filosófico el único que nos puede suministrar las claves centrales de estos conceptos. Pero si diferenciamos los elementos de nuestra imagen manifiesta que podrían ser, o son,

o han sido, revisables o revisados por el estudio científico de la mente de los elementos que no parecen ser revisables sin comprometer ambas imágenes, podríamos plantearnos ahora las cuestiones de la extensión de la mente, ya en el terreno normativo, como una cuestión que se plantea en esos territorios que forman parte de nuestra metafísica supuesta también en la propia construcción científica de nuestra naturaleza. Me parece que las ideas de sujeto, de intencionalidad y de agencia y de identidad personal pertenecen de forma paradigmática a estos territorios. Y entonces podríamos plantear la cuestión de la identidad y extensión de la mente más allá de las fronteras del cuerpo como una cuestión de la identidad de los sujetos, la agencia e intencionalidad y la identidad personal más acá o más allá de las fronteras de lo corporal.

Mi inclinación acerca de los compromisos ontológicos que tienen estos elementos normativos de nuestra auto-conceptualización es bastante liberal: debemos salvar ciertos compromisos normativos que, sin embargo, son bastante neutros ontológicamente. Mi argumento se mueve aquí en el contexto de lo que es el reciente giro práctico hacia la idea de subjetividad que están protagonizando en una línea Richard Moran (2001) y Carol Rovane (1998) y en otros aspectos el inferencialismo de Robert Brandom (1994). El proyecto de Moran es el de cómo salvar los compromisos evaluativos de la idea de lo mental sin caer en el modelo cartesiano de un teatro observado un fantasma que llamamos consciente. Moran propone considerar el acto de expresar un contenido mental en el espacio público como un acto de compromiso en el que el sujeto constituye el estado de su mente en un doble movimiento de autointerpretación y de compromiso público con su afirmación. Es una propuesta sartriana de instaurar el compromiso práctico en el corazón de la racionalidad teórica y práctica. Un caso similar, más complejo en sus consecuencias semánticas es el modelo de Robert Brandom que propone también un sistema de reconocimientos mutuos de compromisos y habilitaciones con las propias expresiones como trasfondo del significado lingüístico. Rovane, por su parte, propone una idea de la identidad personal que considera ética en su raíz: es el reconocimiento de la existencia de un conjunto de proyectos racionales, que sería el componente de identidad narrativa que singulariza a una persona. En estos tres casos, distintos en sus intereses y afirmaciones, subyace una misma posición de sostener las nociones de sujeto, subjetividad e identidad personal sobre los compromisos éticos. No vamos a examinar aquí los defectos y virtudes de estas aproximaciones⁸. Me parece que el núcleo común es hacer de la subjetividad y la identidad personal un resultado derivado de los compromisos prácticos en un espacio público de otras mentes a las que se reconoce como iguales. La idea tiene un elemento normativo, que corresponde a las condiciones de éxito de esos actos constitutivos de los estados mentales propios,

⁸ Me he ocupado de Moran en (Broncano 2006) y de Brandom en (Broncano 2004a).

pero es coherente también con la idea evolucionaria de una mente que ha sido seleccionada en un medio ya social compuesto de otras mentes con las que interactúa.

No estamos diciendo ahora que la mente se extienda, desde el punto de vista normativo, al espacio público, sino más bien que los elementos normativos no habrían devenido en tales sin la existencia de este espacio público. Al constituir la idea de yo, de agencia y de expresión intencional, el agente se autointerpreta en el espejo de la mirada de los otros con los que adquiere compromisos que crean responsabilidad en él. Las ideas de subjetividad y de identidad personal se asientan ahora en ese sistema de compromisos y capacidades de interacción. La autoconciencia y la integración psicológica de la que hemos hablado en las anteriores secciones se incorporan ahora a un nuevo movimiento de afirmaciones y representaciones públicas. Pensemos pongamos por caso la conciencia de la propiedad del propio cuerpo en este espacio complejo: ya no incluye solamente los elementos sensoriales y perceptivos, puede incluir también, por ejemplo, el derecho a no ser observado y preservar la intimidad, el derecho a un cierto espacio propio, etc. Son derechos que progresivamente incorporamos a nuestra identidad y que no son necesariamente correspondientes a las fronteras de lo funcional psicológico o fisiológico. Las personas, así, son en cierto modo conjuntos estables que solamente existen si existe más de una persona.

La mente extendida queda así como una tesis sobre la apertura al mundo de nuestra identidad mental, sobre la hibridación entre lo natural y lo artificial de nuestras funciones psicológicas que no solo no es compatible, sino que encaja de forma mucho más fácil en los entresijos de la fundamentación práctica de nuestra naturaleza.

Para terminar, ¿qué queda de nuestra afirmación inicial de que el funcionalismo significaba una revisión metafísica sustancial? Queda, me parece, como una afirmación que está limitada al terreno de la clasificación ontológica de lo real. No es mucho y no es poco: el funcionalismo ha borrado la frontera entre lo natural y lo artificial; ha borrado la frontera entre lo constitutivo y lo genético; al menos en las formas de funcionalismo que hemos examinado; ha borrado también las fronteras entre lo existente en el espacio y lo existente en el territorio de la abstracción. Mi argumento es que no ha borrado, como tal funcionalismo, las fronteras entre lo descriptivo y lo normativo. Muchos se alegrarán: los filósofos han reclamado para sí este territorio de lo normativo. Bien, nada que objetar, en él estamos los filósofos y de ello comemos; pero, *caveat canem*, esa dicotomía no está tan segura ya como parece y no por culpa ciertamente del funcionalismo, sino de otros cambios, como el llamado “giro de las prácticas”, que ha ocurrido en el siglo XX paralelo al funcionalismo. Pero esa es otra historia.

Referencias bibliográficas

- ADAMS, F. and K. AIZAWA (2001): "The bounds of cognition", *Philosophical Psychology*, 14: pp. 43-64.
- BRANDOM, R. (1994): *Making It Explicit*. Cambridge MA: Harvard University Press
- BRONCANO, F. (2004a): "Una (hegeliana) vuelta de tuerca. Recensión de Robert Brandom", (*Making it Explicit. La articulación de las razones. Tales of the Mighty Dead.*) *Revista de Libros*, Enero (2004), pp. 19-22.
- BRONCANO, F. (2004): "Capacidades metarrepresentacionales y conducta simbólica" *Estudios de Psicología* 25/2, pp. 183-203.
- BRONCANO, F. (2006): "Achieving to make up your mind" (Manuscrito).
- CLARK A. (1997): *Being there*. Cambridge, MA: MIT Press.
- CLARK A. (2006): "Memento's Revenge: The Extended Mind, Extended" URL=www.cogs.indiana.edu/andy/Mementosrevenge2.pdf
- CLARK, A. (2001): *Mindware: An Introduction to the Philosophy of Cognitive Science*. Oxford: Oxford University Press.
- CLARK, A. (2005): "Intrinsic content, Active Memory and the Extended Mind" *Analysis* 65, pp. 1-11.
- CLARK, A., CHALMERS, D. (1998): "The extended mind", *Analysis* 58, pp. 7-19.
- DAMASIO, A.R. (2000): *The Feeling of What Happens: Body, Emotion and the Making of Consciousness* Londres: Vintage.
- DENNETT, D. (1991): *Consciousness Explained*. Boston: Little Brown and Co.. Citado por la trad de Sergio Balari, Paidós.
- DENNETT, D. (1996): *Kinds of Minds*, Londres: Weinfeld & Nicolson.
- DENNETT, D. (1996): *Kinds of minds*. New York, NY: Basic Books.
- DENNETT, D. (2000): "It's Not a Bug, It's a Feature" (Respuesta a Humprey (2000), *Journal of Consciousness Studies* 7/4, pp 25-27.
- DONALD, M. (1991): *Origins of the modern mind*. Cambridge, MA: Harvard University Press.
- DRETSKE, F. (1981): *Knowledge and the Flow of information*. Cambridge, MA: MIT Press.
- DRETSKE, F. (1988): *Explaining behavior*. Cambridge, MA: MIT Press.
- EILAN, N. (2000): "Primitive Consciousness and the 'Hard Problem'" *Journal of Consciousness Studies* 7/4, pp. 28-39.
- HUMPHREY, N. (2000): "How to Solve the Mind-Body Problem" *Journal of Consciousness Studies* 7/4, pp. 5-20.
- HURLEY, S. (1998): "Vehicles, contents, conceptual structure and externalism". *Análisis* 58: pp. 1-6.
- HUTCHINS, E. (1995): *Cognition in the wild*. Cambridge, MA: MIT Press.
- MILLIKAN, R. (1984): *Language, thought, and other biological categories*. Cambridge, MA: MIT Press.

- MORAN, R. (2001): *Authority and Estrangement*, Princeton: Princeton University Press.
- ROVANE, C. (1999): *The Bounds of Agency. An Essay in Revisiorary Metaphysics*. Princeton: Princeton University Press.
- SEARLE, J. (1980): Minds, brains, and programs. *Behavioral and Brain Sciences*, 3, pp. 417-458.
- VAN GELDER, T. & PORT, R. (1995): It's about time: an overview of the dynamical approach to cognition. In R. Port & T. van Gelder (Eds). *Mind as motion: explorations in the dynamics of cognition*. Cambridge, MA: MIT Press.
- VEGA, J. (2006): "Mentes híbridadas: cognición, representaciones externas y artefactos epistémicos" en AIBR. *Revista de Antropología Iberoamericana* / www.aibr.org

Fernando Broncano
Departamento de Humanidades
Universidad Carlos III de Madrid
fernando.broncano@uc3m.es